

## Chapter One: Contents

(Population – 15 October 2001 – LA-UR 01-5714 – Portland Study Reports)

<b>1. GENERAL RESULTS .....</b>	<b>1</b>
1.1 AGREEMENT WITH FORECAST .....	1
1.2 EXAMPLE BLOCK GROUP .....	2
<b>2. SUMMARY .....</b>	<b>6</b>

# Chapter One—Population

## 1. GENERAL RESULTS

### 1.1 Agreement with Forecast

The synthetic population covers 1107 block groups in the greater Portland area. There are a total of 636,389 households, compared to 636,533 households in Portland Metro's forecast (a difference of 144 households). Over 60% of the block groups have exactly the population specified in the forecast. The marginal counts of households for the size, age, and income demographics match very closely: 10.9% of the counts have a  $p$ -level below 10%, 0.99% of the counts have a  $p$ -level below 1%, and 0.21% of the counts have a  $p$ -level below 0.1% for example, if we check the synthetic population with a chi-squared test. There are no differences in the structural zeros for the two sets of marginal distributions.

For a more stringent check of the agreement of the synthetic population with the forecast, we have examined the odds ratio,

$$j = \frac{p_{1,1} \cdot p_{2,2}}{p_{1,2} \cdot p_{2,1}},$$

for randomly-chosen cells in the multiway table for persons, age, and income. (The quantities  $p_{1,1}$ ,  $p_{2,2}$ ,  $p_{1,2}$ , and  $p_{2,1}$  represent the counts in related cells of the multiway table.) Since the iterative proportional fitting procedure preserves the correlation structure (and hence the odds ratio) of the data it fits, there should be a strong correlation between the odds ratios measured in the forecast joint distribution computed by Portland Metro and the synthetic population generated in TRANSIMS. Note that the forecast joint distribution is not used by the Population Synthesizer, so this check constitutes an independent test of the synthesis procedure. Fig. 1 shows a comparison of the cells of the multiway table, while Fig. 2 shows a comparison of the odds ratios—in each case we have good agreement between the forecast and the synthesis.

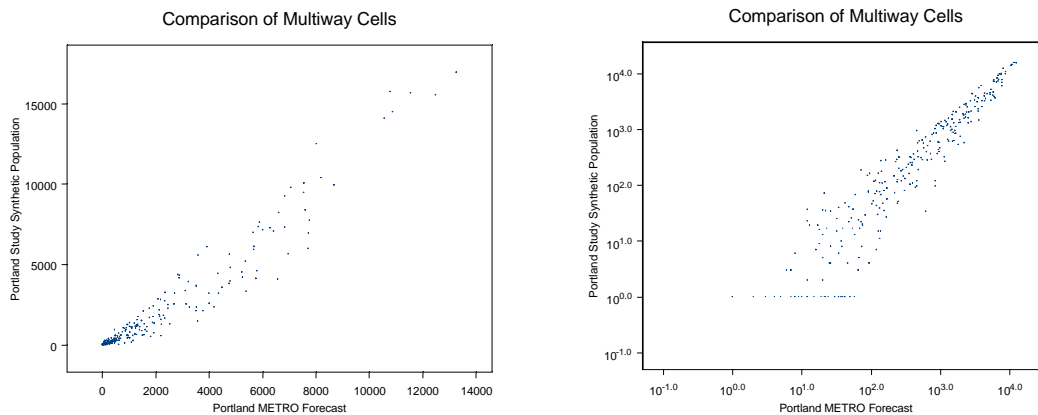


Fig. 1. Comparison of counts in the 441-cell persons  $\times$  age  $\times$  income multiway table representing the correlation structures of the Portland Metro forecast and the Portland study synthetic population: linear scale (left) and logarithmic scale (right). The strong correlation indicates good agreement.

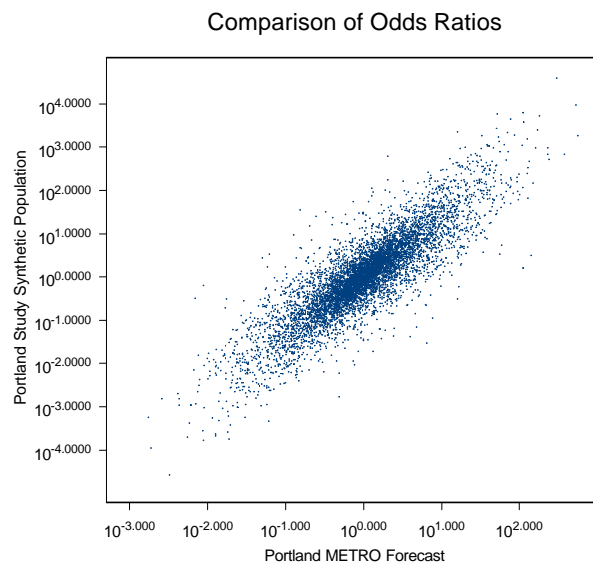


Fig. 2. Comparison of 21,898 randomly-chosen odds ratios for the 441-cell persons  $\times$  age  $\times$  income multiway table representing the correlation structures of the Portland Metro forecast and the Portland study synthetic population. The strong ellipsoidal shape indicates good agreement.

## 1.2 Example Block Group

We choose block group 31200.1 (shown in Fig. 3) to illustrate the results of synthesizing and locating a population; Table 1 lists the forecast marginal demographic data provided by Portland Metro. Fig. 4 shows how the correlation structure of demographic variables is preserved after iterative proportional fitting, while Fig. 5 demonstrates that the marginal distributions of the demographic variables closely match the forecast.

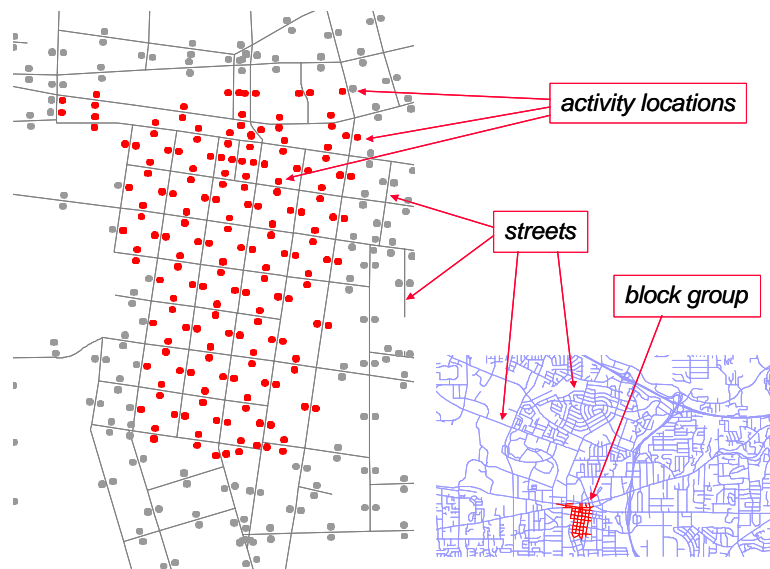


Fig. 3. Geographic layout of block group 31200.1 in Portland, Oregon.

**Table 1. Forecast (1996) marginal demographics for block group 31200.1 in Portland, Oregon.**

Size	Households	Age of Head	Households	Income	Households
1	84	<b>b</b> 24	24	<b>b</b> 4999	0
2	42	25–34	42	5000–9999	47
3	0	35–44	35	10,000–14,999	8
4	6	45–54	28	15,000–24,999	19
5	0	55–64	0	25,000–34,999	37
6	0	65–74	2	35,000–49,999	0
<b>r</b> 7	0	<b>r</b> 75	1	50,000–74,999	21
Total	132	Total	132	75,000–99,999	0
				<b>r</b> 100,000	0
					132

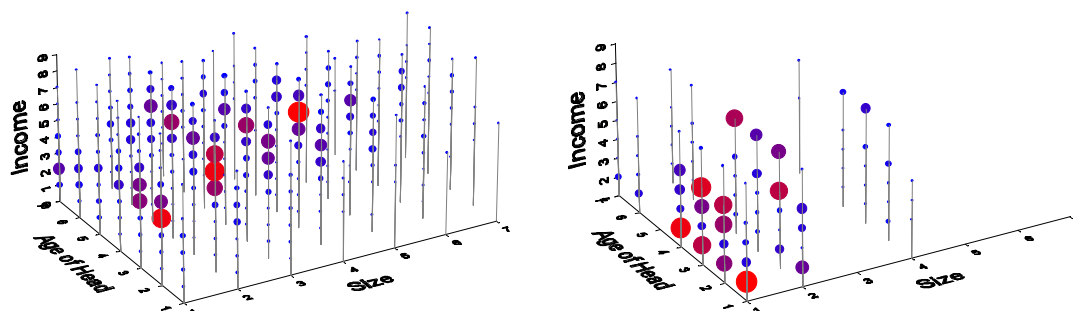
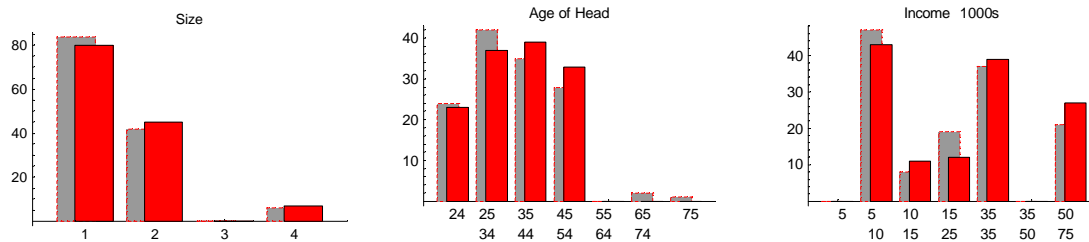


Fig. 4. Comparison of multiway tables for the whole study area (left) and for block group 31200.1 (right) in Portland, Oregon.



*Fig. 5. Comparison between forecast marginal distributions (gray) and counts of households in the synthetic population (red) for block group 31200.1 in Portland, Oregon.*

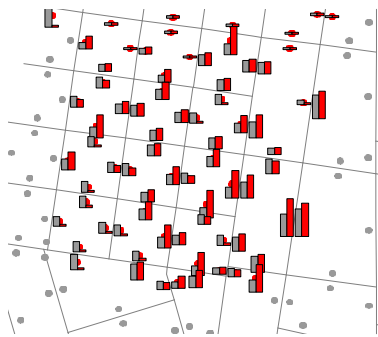
In Table 2 and Table 3, we list the demographic variables generated for an example household from block group 31200.1. Using the population synthesizer, it is possible to generate populations with any of 65 household variables and any of 74 person variables. The variables listed in the tables are those used in the Portland study. Finally, we show in Fig. 6 that the count of households located along streets in block group 31200.1 closely matches the residence area obtained from tax lot data.

**Table 2. Selected person demographics for a synthetic household in block group 31200.1 in Portland, Oregon.**

ID	111733
Size	4
Vehicles	3
Activity Location	23101
PUMS Record	44789
Anyone under 18	Yes
Workers in 1989	3+
Total Income	\$64,000
Tenure	Owned with mortgage or loan
Value	\$90,000 - \$99,999

**Table 3. Selected person demographics for a synthetic household in block group 31200.1 in Portland, Oregon.**

ID	255552	255553	255554	255555
Age	42	42	19	7
Relationship	Householder	Husband/wife	Son/daughter	Son/daughter
Sex	Male	Female	Female	Female
Worked in 1989	Yes	Yes	Yes	No (under 18)
Educational Attainment	Some college, but no degree	High school graduate, diploma or GED	Some college, but no degree	1 <sup>st</sup> , 2 <sup>nd</sup> , 3 <sup>rd</sup> , or 4 <sup>th</sup> grade
Industry	Electrical Machinery, Equipment, and Supplies, N.E.C	Not Specified Retail Trade	Offices and Clinics of Chiropractors	
Occupation	Managers and Administrators, N.E.C	Sales Workers, Other Commodities	Managers, Medicine and Health	
Total Income	\$45,000	\$13,000	\$6000	
Hours Worked	40	40	15	
Lived Here in 1985	No	No	No	(under 5)
Means of transportation to work	Car, truck, or van	Car, truck, or van	Car, truck, or van	
Vehicle occupancy	1	1	1	
Time of departure for work	6:50	1:00	14:00	
Travel time to work	0:20	0:15	0:10	

*Fig. 6. Residence area (gray bars) and count of synthetic households (red bars) for part of block group 31200.1 in Portland, Oregon.*

## 2. SUMMARY

In conclusion, we have generated a statistically realistic synthetic population for use in the Portland study. The synthetic population preserves the key features of the actual population in that it matches its correlation structure, the joint distributions of demographic variables, the spatial distribution of households, and the number of vehicles for households. We have tested to output data from the Population Synthesizer to verify that it performs the necessary calculations correctly and that the results statistically “match” the required constraints. The synthesized population possess a rich and detailed selection of demographic information for individual households and persons.

We also note that the Population Synthesizer has practical input data requirements. It uses readily-available census data (STF-3A, PUMS, TIGER) along with data typically available at metropolitan planning organizations (block-group-level marginal forecasts of population demographics, tax lot and zoning information, and motor vehicle records).